

La lutte contre la fraude à l'ère digitale

Les enjeux du recours à l'intelligence artificielle par l'administration financière

Caroline LEQUESNE-ROTH,

*maître de conférences à l'université Côte d'Azur,
directrice du Master II – Droit algorithmique
et gouvernance des données*



1 - S'il est un domaine où les réalités administratives d'aujourd'hui revêtent les formes des dystopies d'hier, l'administration financière fait figure de modèle. Le « Big Brother » fiscal, et ses tentacules – qui se déploient sur le terrain de l'évaluation financière et du contrôle de légalité¹ –, s'inscrivent désormais dans le paysage de la lutte contre la fraude². Initié dès 2014 à titre expérimental³, le recours à des dispositifs algorithmiques au service de la lutte contre la fraude a été pérennisé en 2017 pour les entreprises, étendu la même année aux contrôles des particuliers⁴. Ces dispositifs recourent aux outils de « *data mining* » – ou algorithmes auto-apprenants –, afin d'automatiser le « ciblage de la fraude et valorisation des requêtes » (CFVR). L'opération consiste à « identifier [...] les critères caracté-

sant une personne fraudeuse afin d'établir un profil type qui sera ensuite appliqué à une population cible »⁵. En d'autres termes, l'administration fiscale recourt aux traitements automatiques des données afin de déterminer des « *patterns* » de fraudes – et leurs éventuels déploiements en réseau – puis les appliquer à des publics cibles.

Le recours à ces outils est plébiscité par l'administration. Les échanges à l'occasion journée d'étude du 28 octobre 2020, organisé sous l'égide de la Société française de finances publiques, en témoignèrent : la Direction générale des finances publiques de la région PACA, qui fit état de ses pratiques, manifesta un enthousiasme certain pour cette assistance, devenue indispensable selon elle. De surcroît, la stratégie gouvernementale de « digitalisation » des contrôles ne semble nullement s'essouffler. Au regard des objectifs affichés tout d'abord : « créer un service national d'analyse de données dont les travaux seraient à l'origine d'environ 60 % des opérations de contrôle fiscal d'ici 2021 »⁶ ; les moyens financiers investis à cet effet depuis 2017 sont conséquents⁷. Du déploiement de nouveaux outils⁸, et de l'élargissement des bases de données au service de la mission CFRV, ensuite : la loi n° 2019-1479 du 28 décembre 2019 de finances pour 2020 a en effet étendu cette base aux réseaux sociaux, permettant le « *scrapping* » de données sur ces derniers⁹. Du lancement du projet PILAT (projet de pilotage et d'analyse du contrôle)

1. V. en ce sens, A. NOR : CPAE1903196A, 29 janv. 2019, portant création d'un traitement automatisé d'analyse prédictive relatif au contrôle de la dépense de l'État. Ce dispositif a vocation à assister les comptables publics dans le cadre de certaines vérifications : qualité de l'ordonnateur, disponibilité des crédits, production des pièces justificatives, etc. Mis en place à titre expérimental au sein de la Direction régionale des finances publiques de Bretagne en 2019, il est appelé à se généraliser à l'horizon 2020-2021.
2. En mars 2019, le président de l'Assemblée nationale a ainsi annoncé, parmi ses propositions visant à « rénover la vie parlementaire », la création d'un nouvel outil « LexImpact » dont l'objet est de permettre, à terme, « tout parlementaire voulant chiffrer un amendement [...] [de] lui-même s'en charger grâce aux logiciels », *Éric Woerth, Séance en hémicycle, mercredi 19 juin 2019*.
3. Les premières expérimentations visaient les entreprises, A. NOR : BUDE1405018A, 21 févr. 2014, portant création par la direction générale des finances publiques d'un traitement automatisé de lutte contre la fraude dénommé « ciblage de la fraude et valorisation des requêtes ». En 2017, elles ont été élargies aux particuliers, *Avis CNIL, délib. n° 2016[shy] 286 ; A. NOR : CPAE1728989A, 28 août 2017, modifiant l'arrêté du 21 février 2014 portant création par la direction générale des finances publiques d'un traitement automatisé de lutte contre la fraude dénommé « ciblage de la fraude et valorisation des requêtes »*.
4. Le dispositif pilote visant à améliorer le ciblage des entreprises fraudeuses a été pérennisé en 2017 ; le dispositif concernant les particuliers doit encore faire l'objet d'une évaluation, conformément aux prescriptions de la CNIL.

5. *Rapp. Sénat n° 668, 22 juill. 2020.*

6. *Rapp. Sénat n° 668, 22 juill. 2020, p. 54.*

7. *Rapp. Sénat n° 668, 22 juill. 2020, p. 54.*

8. La DGFIP a ainsi commencé le développement du « *text mining* », un traitement de données non structurées, pour les déclarations de succession et sur les actes de cessions de biens immobiliers, *Rapp. Sénat n° 668, 22 juill. 2020, p. 58.*

9. L'article 154 de la loi dispose qu'« [à] titre expérimental et pour une durée de trois ans, pour les besoins de la recherche de [certains] manquements et infractions, [...] l'administration fiscale et l'administration des douanes et droits indirects peuvent, chacune pour ce qui la concerne, collecter et exploiter au moyen de traitements informatisés et automatisés [...] librement accessibles sur les sites internet des opérateurs de plateforme en ligne [...] manifestement

enfin, dont l'ambition est de « supprimer les ruptures applicatives sur l'ensemble de la chaîne du contrôle fiscal, moderniser et simplifier le travail de l'agent participant à la chaîne du contrôle et améliorer le pilotage de la mission et la valorisation de l'activité »¹⁰.

Cette volonté gouvernementale est mue par diverses considérations ; volumétriques en premier lieu : la mission requête et valorisation de la Direction générale des finances publiques (ci-après DGFIP) traite plus de 200 téraoctets (To) de données issues des échanges automatiques d'informations réceptionnées par l'application EAI, et des données en provenance de 92 pays¹¹ ; en 2019, l'administration fiscale a reçu des informations relatives à 4,8 millions de comptes à l'étranger détenus par des contribuables français. La dynamique répond également, et en deuxième lieu, de l'exigence de « performance » en termes de rendement des contrôles : le *data mining* doit renforcer l'efficacité du traitement¹². Elle est en troisième lieu animée, bien que de manière implicite, par le souci constant de réduction des coûts et du nombre de fonctionnaires. Elle s'inscrit en quatrième lieu dans une stratégie globale, résultant de la libéralisation et de l'ouverture des frontières. Si ces phénomènes ont engendré de nouveaux défis en matière de lutte contre la fraude (évasion fiscale, forum shopping), ils ont également conduit à renforcer les collaborations bilatérales, l'échange d'informations et la standardisation des pratiques administratives que la technologie accélère.

Notons enfin que la France n'est pas isolée dans sa démarche : l'automatisation – du moins partielle – de la lutte contre la fraude fiscale connaît des développements similaires en Grande-Bretagne (au travers du dispositif « Connect »), en Italie (avec l'indicateur « redditometro ») ou encore en Belgique, dont l'administration fut pionnière dans le développement d'outils algorithmiques de lutte contre les carrousels de TVA.

La journée de travail du 28 octobre, et les présents travaux qui la poursuivent, avaient pour ambition d'appréhender le phénomène de digitalisation dans ses divers aspects : son encadrement juridique¹³, les enjeux de la traduction algorithmique¹⁴ et l'avenir du contrôle fiscal considéré depuis l'entreprise¹⁵. Il ressort de nos échanges, un constat : celui d'une transformation, profonde et pérenne, du contrôle fiscal français dans ses formes (1), sa nature (2) et du paradigme dans lequel il s'insère, consacrant l'avènement d'une société de surveillance (3).

1. Transformation dans des formes du contrôle

2 - L'introduction d'outils de ciblage, et plus largement d'une assistance algorithmique, dans le processus de contrôle, invite à repenser le rôle de l'administration. La dynamique engagée entend positionner principalement celle-ci sur le volet des vérifications et des

sanctions, le repérage, l'initiative et l'opportunité des contrôles étant dévolus au dispositif technique. Bien que cette répartition puisse en pratique s'apparenter à une délégation – du moins partielle – des fonctions de l'administration, le Conseil constitutionnel a réfuté cette lecture. Selon lui, l'exercice constitutionnel du pouvoir réglementaire n'est nullement en cause dès lors que les dispositifs algorithmiques « n'ont ni pour objet ni pour effet d'autoriser l'administration à adopter des décisions sans base légale, ni à appliquer d'autres règles que celles du droit en vigueur »¹⁶. Cette lecture, contestable sur le fond, l'est encore sur la forme : l'absence de délégation de la part de l'administration suppose une pleine maîtrise par celle-ci de la procédure et des étapes du processus décisionnel. Or, tel n'est manifestement pas le cas de l'administration française dont « l'environnement d'exploitation numérique » est jugé défaillant. En cause : l'insuffisante compétence des agents, l'absence de politique générale de formation¹⁷ ou encore le défaut de « véritable management » algorithmique¹⁸, en amont et en aval de l'adoption des dispositifs, dans les services in house comme dans les rapports avec les prestataires techniques.

Le contrôle humain se mue ainsi, *de facto*, en contrôle hybride, pour lequel notre droit, conçu pour la décision proprement humaine, n'est pas adapté. Cela est d'autant plus problématique¹⁹ que la technique n'est pas neutre, susceptible d'erreurs et sujette à engendrer des discriminations. L'administration britannique en offre un exemple significatif : en septembre 2020, le gouvernement indiquait que l'adoption d'un dispositif algorithmique défaillant avait engendré une perte de recettes évaluée à 31 milliards de livres sterling. Le dispositif, conçu pour le système fiscal américain, s'était révélé inadapté au calcul de l'écart fiscal britannique²⁰. Comment l'administration peut-elle identifier et corriger les biais et les erreurs algorithmiques sans en avoir la maîtrise ni même la compréhension élémentaire ?

2. Transformation de la nature du contrôle

3 - La transformation du contrôle fiscal peut encore être appréhendée dans sa nature : ces dispositifs contribuent à ériger le fait en droit, entraînant la migration d'un système juridique causal vers un système de corrélations pratiques.

En effet, le fraudeur n'est pas tant ciblé à partir de l'infraction commise qu'au départ de signaux, comportementaux et/ou relationnels, identifiés itérativement par le dispositif algorithmique. La détection des fraudeurs résulte, dans ce contexte, non pas d'une suspicion fondée sur des présomptions, mais de probabilités statistiques « auto-engendrées » par le dispositif. Les faits, et les correspondances factuelles, permettent d'établir des profils types (ou *pattern*), qui opèrent comme des grilles d'évaluation : ils fondent l'attribution des notes (*scoring*) évaluant le risque de fraude. La note conditionne dans une large mesure les poursuites, révélant ainsi sa nature normative. Dans ce schéma, l'opportunité des contrôles repose sur une norme fiscale algorithmique, résultant d'une analyse proprement

rendus publics par leurs utilisateurs », V. L. n° 2019-1479, 28 déc. 2019, art. 154 : Dr. fisc. 2020, n° 1, comm. 51.

10. Lutte contre l'évasion fiscale et la fraude en matière d'impositions de toutes natures et de cotisations sociales : Doc. de politique transversale projet de loi de finances pour 2021, déc. 2020, p. 40.

11. Lutte contre l'évasion fiscale et la fraude en matière d'impositions de toutes natures et de cotisations sociales : Doc. de politique transversale projet de loi de finances pour 2021, déc. 2020, p. 40.

12. Notons pour autant que la rhétorique gouvernementale s'accorde mal des premiers bilans qui pour l'heure ne font pas la démonstration d'une rentabilité des contrôles accrue, ni d'économie d'échelle, V. en ce sens, *concl. Rapp. Sénat* n° 668, 22 juill. 2020.

13. P. Legros et M. Kimri, *Le régime juridique du contrôle fiscal algorithmique* : Dr. fisc. 2021, n° 5, étude 122.

14. D. Merigoux et L. Huttner, *Traduire la loi en code grâce au langage de programmation Catala* : Dr. fisc. 2021, n° 5, étude 121.

15. J. Sayag, *L'entreprise à l'épreuve de l'évolution des outils du contrôle fiscal* : Dr. fisc. 2021, n° 5, étude 123.

16. *Cons. const.*, 12 juin 2018, n° 2018-765 DC sur la loi relative à la protection des données personnelles : *Comm. com. electr.* 2018, comm. 65, obs. A. Debet.

17. *Rapp. Sénat* n° 668, 22 juill. 2020, p. 21.

18. *Rapp. élèves de l'ENA, Éthique et responsabilité des algorithmes publics*, juin 2019, disponible sur <https://www.etalab.gouv.fr/wp-content/uploads/2020/01/Rapport-ENA-Ethique-et-responsabilite%20C3%A9-des-algorithmes-publics.pdf>.

19. Comme nous l'avons démontré par ailleurs, C. Lequesne-Roth, *La science des données numériques au service du contrôle fiscal français. Réflexions sur l'Algocratie*, in A. Pariente (dir.), *Les chiffres en finances publiques : Mare & Martin*, 2019, p. 189 et 190.

20. <https://www.thisismoney.co.uk/money/news/article-8701583/Now-algorithm-gets-UKs-31bn-tax-gap-wrong.html>.

factuelle. Les dispositifs produisent ainsi, au départ des faits, une nouvelle forme de norme, proprement algorithmique.

Cette transformation emporte, par ricochet, une mutation du raisonnement juridique lequel, jusqu'alors « causal », prend la forme d'un système de corrélations pratiques²¹. Il ne s'agit plus tant d'identifier les fraudeurs au départ d'une analyse contextuelle et qualitative, que d'apprécier une probabilité déterminée par des paramètres largement quantitatifs.

Ces transformations soulèvent, de nouveau, des questions préoccupantes sur le terrain de l'état de droit : quelle(s) garantie(s) offre(nt) ces dispositifs – et les normes qu'ils génèrent – quant à leur conformité au droit fiscal ? Ces normes techniques, dont la légitimité politique fait défaut, sont-elles conformes à nos droits fondamentaux ?

3. Transformation d'un paradigme

4 - L'introduction de dispositifs algorithmiques suscite, plus fondamentalement, un changement de paradigme. Le terme de « Big Brother » fiscal, déjà mentionné, et ancré dans la culture populaire, décrit ce constat : la digitalisation de la lutte contre la fraude fiscale participe de l'avènement de l'État de surveillance. Les outils algorithmiques de lutte contre la fraude appartiennent à l'appareillage sécuritaire. Ils impliquent, dans un même mouvement, une surveillance *ex ante* – la fraude étant présumée avant même sa réalisation –, et la surveillance généralisée de l'ensemble des individus pour repérer les comportements frauduleux de quelques-uns. Le recours annoncé – et permis – au *web scrapping* en constitue une illustration : il conduit à renverser la présomption de culpabilité en consacrant le contrôle de l'ensemble des citoyens.

Cette tendance, dénoncée par le rapporteur spécial aux droits de l'homme de l'ONU, constitue un dévoiement inquiétant de l'État providence : « Le processus communément appelé “transformation numérique” par les gouvernements [...] ne doit pas occulter le caractère révolutionnaire et politiquement motivé de nombreuses innovations de ce type [...] utilisées à des fins diverses, notamment pour

automatiser, prévoir, identifier, surveiller, détecter, cibler et punir »²².

Une pente glissante pourtant jugée proportionnée par le Conseil constitutionnel, dont la décision relative à l'élargissement des bases de données a accordé au législateur un satisfecit, sous quelques modestes réserves : « pour la recherche des manquements et infractions [...] le législateur a assorti le dispositif critiqué de garanties propres à assurer, entre le droit au respect de la vie privée et l'objectif de valeur constitutionnelle de lutte contre la fraude et l'évasion fiscales, une conciliation qui n'est pas déséquilibrée. Il en résulte également que l'atteinte à l'exercice de la liberté d'expression et de communication est nécessaire, adaptée et proportionnée aux objectifs poursuivis »²³. Une décision contestable au regard du précédent qu'elle forme, le Haut Conseil avalisant le déploiement de dispositifs de surveillance, qui se multiplient aujourd'hui dans la plupart de nos administrations²⁴.

L'évolution du contrôle fiscal dans ses formes digitales se présente ainsi, à bien des égards, comme un indicateur : des transformations de la normativité à l'ère de l'administration algorithmique ; des transformations, encore, de nos régimes démocratiques, dont elle décrit une dynamique préoccupante. À l'avenir, un développement de ces outils efficace et respectueux de l'état de droit supposera aussi une compréhension de leur contexte de déploiement (environnement numérique adapté), et des mécaniques de leur fonctionnement. Il supposera également de concevoir le contrôle de ces dispositifs, qui pourrait prendre la forme de « chaînes d'audits », ouverts à des tiers habilités, en amont comme en aval de leur adoption. En tout état de cause, nous plaçons en faveur d'un contrôle politique, pluriel et technique. La journée d'étude du 28 octobre en a rappelé l'importance, tout comme celle des débats quant à l'opportunité des usages de la technologie au sein de l'administration fiscale, débats quasi inexistant à ce jour au regard de la supposée « neutralité » technologique. Il en va ainsi de la démocratisation du contrôle fiscal algorithmique.

21. D. Cardon, *A quoi rêvent les algorithmes : Seuil*, 2015, p. 39 à 44. – A. Garapon et J. Lassègue, *Justice digitale, Révolution graphique et rupture anthropologique* : PUF, 2018, p. 223.

22. *Report of the Special rapporteur on extreme poverty and human rights*, A/74/48037, 11 oct. 2019.

23. *Cons. const.*, 27 déc. 2019, n° 2019-796 DC : *Dr. fisc.* 2020, n° 1, comm. 6.

24. Des organismes sociaux à la police. Le rôle moteur de l'administration fiscale a en outre été rappelé dans le rapport du Sénat, lequel indique en ce sens que « ses usages et techniques [sont] susceptibles d'être reproduits dans d'autres sphères que celles du contrôle fiscal », V. *Rapp. Sénat* n° 668, 22 juill. 2020, p. 56.